REMARKS

Applicants have carefully reviewed the non-final Office Action mailed December 27, 2007, and thank Examiner Siedler for the detailed review of the pending claims. In response to the Office Action, Applicants have amended claims 2, 19, 24, and 31, and added claim 32. By way of these amendments, no new matter has been added. More specifically, the amendments are supported at least at paragraphs [0052] – [0056] of the specification, and FIGS. 4 and 5 of the drawings as originally filed. Claim 13 was canceled previously. Accordingly, claims 1-12 and 14-32 remain pending in this application.

As Applicant's remarks with respect to the Examiner's rejections are sufficient to overcome these rejections, Applicant's silence as to assertions by the Examiner in the Office Action or certain requirements that may be applicable to such rejections (e.g., whether a reference constitutes prior art, motivation to combine references, assertions as to dependent claims, etc.) is not a concession by Applicant that such assertions are accurate or such requirements have been met, and Applicant reserves the right to analyze and dispute such assertions/requirements in the future. Further, for any instances in which the Examiner took Official Notice in the Office Action, Applicants expressly do not acquiesce to the taking of Official Notice, and respectfully request that the Examiner provide an affidavit to support the Official Notice taken in the next Office Action, as required by 37 CFR 1.104(d)(2) and MPEP § 2144.03.

Rejections Pursuant to 35 U.S.C. §103

Claims 1-12, and 14-31 were rejected under 35 U.S.C. §103(a) as being allegedly unpatentable over Colbath ("Spoken Document: Creating Searchable Archives from Continuous Audio" IEE 2000) in view of Liu ("Fast Speaker Change Detection for Broadcast News Transcription and Indexing", and further in view of Stanford (U.S. Patent No. 5,475,792). These rejections are respectfully traversed.

A. Obviousness

With respect to the present Section 103 rejections, Applicants respectfully submit that the Office Action has failed to meet the burden of stating a *prima facie* case of obviousness. A *prima facie* case of obviousness requires that:

> first, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations.

See MPEP, § 2143 (citing *In re Vaeck*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991). Applicant further submits that the first requirement for a *prima facie* case of obviousness set forth above is not inconsistent with the United States Supreme Court's recent decision in *KSR International Co. v. Teleflex, Inc.*, ___ U.S. ___, (April 30, 2007) (citations herein are taken from the Court's Bench Opinion). In *KSR*, the Supreme Court stated that "[t]he combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results." Id. at 12. Additionally the court stated:

> it can be important to identify a reason that would have prompted a person of ordinary skill in the relevant field to combine the elements in the way the claimed new invention does. This is so because inventions in most, if not all, instances rely upon building blocks long since uncovered, and claimed discoveries almost of necessity will be combinations of what, in some sense, is already known.

Id. at 15. The Court further explained that

> What matters is the objective reach of the claim. If the claim extends to what is obvious, it is invalid under §103. One of the ways in which a patent's subject matter can be proved obvious is by noting that there existed at the time of invention a known

> problem for which there was an obvious solution encompassed by the patent's claims.

Id. at 16. Accordingly, the Court made clear that "a patent composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known and in the prior art." Id. at 14. In summary, *KSR* plainly does not disturb the well-settled proposition that a reference must be considered in its entirety, i.e., as a whole, including portions that would lead away from the claimed invention. *W.L. Gore & Associates, Inc. v. Garlock, Inc.*, 721 F.2d 1540, 220 USPQ 303 (Fed. Cir. 1983), *cert. denied*, 469 U.S. 851 (1984); M.P.E.P § 2141.02.

B. Independent Claims 1, 10, 16, 23, and 28

Independent claims 1 and 28 recite in part "at least one model for classifying the sound in the audio signal based on bandwidth." Similarly, independent claim 10 recites "training at least one model to classify the sound based on a bandwidth of the sound" and independent claim 16 recites "models that classify the portions of the audio signal based on bandwidth." Further, independent claim 23 recites "audio classification logic configured to classify the input audio data into at least one of a plurality of broad audio classes, the broad audio classes including… a non-phoneme based bandwidth class." One of skill in the art at the time the invention was made would not have combined the references relied on by the Examiner to re-create the claimed arrangements.

Colbath teaches an audio indexing system called "Rough 'n' Ready" that produces a "rough" transcription of audio files that is "ready" to be browsed. *See Colbath, Page 2, 6th paragraph.* The system uses speaker segmentation, clustering, and identification, speech recognition, name spotting, topic classification, and story segmentation. *See Colbath, Page 2, 6th full paragraph.* Specifically, Colbath employs speaker identification and segmentation to "create paragraph-like units between speakers, both known and unknown (classified by gender). Speaker change detection is important for correct playback of audio sections of the archive, and speaker identification allows the user of the archive to perform queries about particular speakers known

to the system, and to skim over the archive for areas where particular speakers were present. In general, however, most speakers will be unknown to the system, so the speaker identification system will cluster them and give them a unique name." (emphasis added) *See Colbath, Page 3, 2nd paragraph.* The Office Action relies upon this passage of Colbath as teaching "classifying the sound in the audio signal based on at least one non-phoneme based model (page 3, Speaker Segmentation, Clustering, and Identification, *the speakers are classified by gender for segmentation and identification*)." (emphasis in original) *See Office Action dated 12/27/2007, page 3.* However, the above-quoted passage of Colbath referenced in the Office Action merely refers to the grouping of speakers for searching in an index, and is not referring to classification of an "audio signal," "sound," or "audio data," much less classification based upon "bandwidth," as recited in the independent claims.

The Office Action acknowledged that Colbath fails to teach or suggest classification according to "bandwidth" of a sound, and combines Liu and Stanford with Colbath in rejecting each of the independent claims. More specifically, the Office Action relies on Stanford as teaching classification "based on bandwidth," stating:

> Stanford discloses a speech recognition system that enables recognition of high bandwidth or telephony (low bandwidth) speech signals (column 2 lines 30-32 and column 8 lines 36-44). Stanford states that low bandwidth speech reduces the accuracy of speech recognizers (column 3 lines 37-39), and discloses a system that trains and uses two separate codebook and phoneme models, one for low bandwidth speech and one for high bandwidth speech (column 8 lines 36-44). The addition of the low bandwidth recognition model improves recognition accuracy for low bandwidth input, such as telephone speech.
>
> Therefore, it would have been obvious to one or (*sic*) ordinary skill in the art at the time of the invention to classifying the sound in the audio signal based on at least one non-phoneme based model, the at least one non-phoneme model including at least one model for classifying the sound in the audio signal based on bandwidth in Colbath, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to achieve the

> predictable results of accurately recognizing and segmenting audio information for indexing, regardless of the input quality.

See Office Action dated 12/27/2007, pages 4-5.

First, the passages of Stanford referenced above merely observe that "bandwidth reductions and noise introduced by telephone lines reduce the accuracy of all speech recognition systems." *See Stanford, Col. 3, lines 37-39.* Stanford solves this accuracy problem with a method that trains a voice recognition system for narrower-bandwidth applications such as telephony by simulating the effects of a telephone system on a sample sound stream. *See Stanford, Col. 12, line 44 – Col. 13, line 11.* Stanford thus merely employs a method for increasing the accuracy of voice recognition systems for narrower-bandwidth applications, and has nothing to do with classification of sounds or, more specifically, "at least one model for classifying the sound in the audio signal based on bandwidth," as recited in part by independent claims 1 and 28, "training at least one model to classify the sound based on a bandwidth of the sound," as recited in part by independent claim 10, "models that classify the portions of the audio signal based on bandwidth," as recited in part by independent claim 16, or "audio classification logic configured to classify the input audio data into at least one of a plurality of broad audio classes, the broad audio classes including… a non-phoneme based bandwidth class," as recited by independent claim 23.

Second, the Examiner's conclusion that "it would have been obvious to one or ordinary skill in the art at the time of the invention to classifying the sound in the audio signal based on at least one non-phoneme based model, the at least one non-phoneme model including at least one model for classifying the sound in the audio signal based on bandwidth in Colbath, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to achieve the predictable results of accurately recognizing and segmenting audio information for indexing, regardless of the input quality" is overbroad, and ignores the purpose of each reference and the effect of the proposed combination on their respective purposes. Merely as an example, the Office Action assumes that Colbath would benefit by employing complex phoneme-based models for the simple task of classifying speakers associated with audio data. However, Colbath

is simply directed to creating a "rough" transcription for searching, and actually sacrifices quality in favor of increased speed in the transcription process. More specifically, Colbath states in regard to the speech recognition component that "we accept a drop in performance to 76% overall word accuracy" in order to "run the system at a faster speed." *See Colbath, Page 3, 1st paragraph.* This is allowable, and even encouraged by Colbath, because absolute quality is not necessary in the transcription in order to properly identify speakers associated with the audio data that is transcribed. Colbath indicates this purpose is by the very name of the system, titling it "Rough 'n' Ready," as it is aimed at producing a "ROUGH transcription, which is READY to be browsed." (emphasis in original) *See Colbath, Page 2, 6th paragraph.* Applicants respectfully submit that the Office Action is selectively ignoring the purpose of the Colbath reference in attempting to combine Colbath with any classification system or method relying on "bandwidth." The only teaching presently before the Examiner that recommends such a combination is that in Applicants' specification.

Further, even if Colbath's purpose could be ignored to allow combining Colbath with Stanford, the Examiner's combination of the references in re-creating the independent claims is specifically discouraged by the teachings of Liu, which is also included in the combination. More specifically, Liu employs a reduced number of phone and non-speech classes to increase decoding speed by approximately 30 times over the base algorithm. *See Liu, Section 5: "Speed."* Further, Liu claims that accuracy is *improved* with the proposed model, underscoring an *advantage* of employing fewer classes for speech classification. *See Office Action dated 10/5/2007, pages 2-3.* In other words, Liu teaches a reduced number of classes for non-speech events, resulting in greater speed *and* accuracy, *thereby providing an additional reason for* ***not*** *relying on additional classes*, e.g., a non-phoneme based class, a bandwidth-based class, or the like, to the model proposed by Liu. In other words, if the model proposed by Liu using a minimal number of phoneme-based classes is not only faster, but also more accurate, there could not possibly be any motivation for adding additional classes, e.g., a bandwidth-based class to Colbath or Stanford, to improve accuracy or speed. Thus, one of ordinary skill in the art at the time the invention was made would not have attempted the combination proposed by the Office

Action for at least this additional reason. Accordingly, independent claims 1, 10, 16, 23, and 28 are allowable over the references of record.

C. Dependent Claims

The dependent claims, by depending from independent claims 1, 10, 16, 23, and 28, are believed to be allowable over the references of record. Moreover, the dependent claims recite independently patentable subject matter. Merely by way of example, the references of record, and in particular Liu, specifically teaches away from classifying sound "based on speaker gender," as recited in claims 2 and 19, and "based on gender of a speaker," as recited in claim 14. As described above in regard to independent claim 23, Liu discourages employing gender-based classification models because they require complicated heuristics to provide adequate robustness. *See Liu, Section 3: paragraph 3, lines 9-15.*

Additionally, claims 2, 19, 24, 31, and 32 each recite, in part, methods and systems for classifying sound or audio signals as belonging to "one of the vowel class, the fricative class, a coughing class, and a silence class," classifying sound or audio signals as belonging to "one of a narrowband class and a wideband class after classifying the sound in the audio signal in the one of the vowel class, the fricative class, the coughing class, and the silence class," and classifying sound or audio signals as belonging to "one of a male class and a female class after classifying the sound in the audio signal in the one of the narrowband class and the wideband class." These recitations are neither taught nor suggested by the cited references. More specifically, there is no teaching whatsoever regarding the order in which sounds or audio data may be classified, or how various classifications may be grouped together in order to classify a sound into a particular category in a particular step of a classification method or system. Accordingly, dependent claims 2, 19, 24, 31, and 32 are allowable over the references of record for at least these additional reasons.

CONCLUSION

All rejections have been addressed. In view of the above amendments and remarks, Applicants believe the pending application is in condition for allowance.

Applicant believes no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 18-0013, under Order No. 65632-0233 from which the undersigned is authorized to draw.

Dated: March 27, 2008

Respectfully submitted,

Electronic signature: /Stefan D. Osterbur/
Michael B. Stewart
Registration No.: 36,018
Stefan D. Osterbur
Registration No.: 57,954
RADER, FISHMAN & GRAUER PLLC
Correspondence Customer Number: 25537
Attorneys for Applicant